

### **Amendments to the Claims:**

This listing of claims will replace all prior versions, and listings, of claims in the application:

### **Listing of Claims:**

1. (Currently Amended) A method, in a data processing system, of distributing traffic to application instances on one or more computing devices, comprising:

obtaining application instance specific operational information identifying operational characteristics of an application instance on a computing device of the one or more computing devices, wherein the application instance specific operational information includes at least one of an application instance topology, an importance of transactions currently being processed by the application instance, an amount of time the application instance has been blocked waiting for resources, and an amount of resources consumed by the application instance;

~~generating a load balancing weight to be associated with the application instance based on the application instance specific operational information obtained; and~~

comparing the application instance specific operational information to one or more other application instance specific operational information for one or more other application instances based on the application instance specific operational information obtained;

generating a load balancing weight based on a relationship between the application instance specific operational information and the one or more other application instance specific operational information;

attributing weight points to the application instance and the one or more other application instances based on a relative difference between the application instance specific operational information and the one or more other application instance specific operational information;

distributing the traffic to the application instance based on the load balancing weight[[]];

wherein obtaining application instance specific operational information includes retrieving the application instance specific operational information from the application instance using an agent application residing on the computing device, and wherein the agent application identifies the application instance topology by sending a correlation in a request to an agent application associated with a second application instance, wherein application instance information is provided by the agent application associated with the second application; and

wherein the method is implemented in a weight management system that is separate from the computing devices and from a load balancing device.

2. (Canceled)
3. (Previously Presented) The method of claim 1, wherein the application instance is instrumented to include code for communicating with the agent application and sending the application instance specific operational information to the agent application from the application instance.
4. (Canceled)
5. (Currently Amended) The method of claim [[4]] 1, wherein the relationship is a relative difference between the application instance specific operational information and the one or more other application instance specific information.
6. (Canceled)
7. (Canceled)
8. (Previously Presented) The method of claim 1, wherein retrieving the application instance specific operational information from the agent application is performed periodically.
9. (Canceled)
10. (Currently Amended) The method of claim [[4]] 1, wherein generating the load balancing weight based on the relationship between the application instance specific operational information and the one of more other application instance specific information includes:
  - assigning a base weight to each of the application instance and the one or more other application instances; and
  - increasing a weight value associated with the application instance or the one or more other application instances based on one or more of the following:
    - which of the application instance and the one or more other application instances has a relatively higher transaction success rate;
    - which of the application instance and the one or more other application instances operates on an underutilized system;
    - ~~which of the application instance and the one or more other application instances has a relatively better response time and operates on an underutilized system;~~

which of the application instance and the one or more other application instances processes the least significant transactions; and

which of the application instance and the one or more other application instances passes transactions on to higher performing computing systems.

11-20. (Canceled)

21. (Currently Amended) The method of claim [[4]] 1, wherein generating the load balancing weight based on the relationship between the application instance specific operational information and the one of more other application instance specific information includes:

assigning a base weight to each of the application instance and the one or more other application instances; and

increasing a weight value associated with the application instance or the one or more other application instances based on which of the application instance and the one or more other application instances processes the least significant transactions.

22. (Currently Amended) The method of claim [[4]] 1, wherein generating the load balancing weight based on the relationship between the application instance specific operational information and the one of more other application instance specific information includes:

assigning a base weight to each of the application instance and the one or more other application instances; and

increasing a weight value associated with the application instance or the one or more other application instances based on which of the application instance and the one or more other application instances passes transactions on to higher performing computing systems.